Vision Research 49 (2009) 1363–1377

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Attention and the detection of masked radial frequency patterns: Data and model

Philip L. Smith [a,*], Young-Eun Lee [a], Bradley J. Wolfgang [a], Roger Ratcliff [b]

[a] Department of Psychology, The University of Melbourne, Vic. 3010, Australia
[b] The Ohio State University, USA

ARTICLE INFO

Article history:
Received 11 October 2007
Received in revised form 21 April 2008

Keywords:
Attention
Spatial cuing
Visual masking
Radial frequency masking
Shunting equation

ABSTRACT

A radial frequency (RF) stimulus is strongly masked by a second, surrounding RF stimulus that follows the first after a critical stimulus onset asynchrony (SOA) of around 100 ms. We sought to determine whether a mask-dependent attentional cuing effect, like that found when detecting pattern-masked sinusoidal gratings, would be obtained with RF stimuli. Observers detected RF modulations in cued or miscued stimuli that were masked with a simultaneous (SIM) RF mask or a delayed (SUC) RF mask that followed it after 100 ms. There were large cuing effects in the SUC condition and small cuing effects in the SIM condition, replicating previous findings. The data are well described by a model in which masks affect the informational persistence of stimuli and cues affect the rate at which stimulus information is transferred into visual short-term memory.

1. Introduction

In this article, we investigate the relationship between attention and a new form of visual masking, *radial frequency pattern masking*, reported by Habak, Wilkinson, Zakar, and Wilson (2004) and Habak, Wilkinson, and Wilson (2006). This form of masking is found with pairs of radial frequency (RF) pattern stimuli, like those depicted in Fig. 1. An RF stimulus is formed by sinusoidally modulating the radius of a suprathreshold-contrast circle that is band-limited in spatial frequency. The observer's task is to judge whether the stimulus that is presented is a true circle (zero RF modulation), or has been radially deformed. The measure of interest is the modulation threshold, that is, the minimum depth of modulation needed to distinguish between a deformed and an undeformed circle. Such stimuli are interesting theoretically because they appear to be processed holistically—at least at low modulation frequencies (Bell, Badcock, Wilson, & Wilkinson, 2007). This has been taken as evidence that they stimulate higher-order visual mechanisms, possibly in area V4, which may act as basis functions for human pattern perception (Habak et al., 2004).

Habak et al. (2006) showed that an RF stimulus is strongly masked by a second RF stimulus whose contours surround the first stimulus, but do not touch or overlap it, and which follows the first after a critical delay of 80–110 ms. The masking function that is obtained is a strongly U-shaped, or Type B, masking function, of the kind that has been found in metacontrast masking and masking-by-structure paradigms (Breitmeyer, 1984; Breitmeyer & Ogmen, 2000). There is relatively little masking when the mask precedes the target (forward masking) or when the target and mask are presented simultaneously. When the mask trails the target (backward masking), the magnitude of the masking effect increases sharply with increasing mask delay up to the critical target-mask stimulus onset asynchrony (SOA) and then decreases. An RF mask of the same frequency and phase angle as the target, presented at the critical SOA, can produce a 16-fold elevation of modulation thresholds. Fig. 2 shows a masking function of this kind. There is a relatively small amount of forward masking, but substantial backward masking.

In the masking literature, this type of masking function has been distinguished from a second, V-shaped, or Type A masking function (Breitmeyer, 1984). Unlike Type B functions, Type A masking functions are symmetrical; forward and backward masking effects are equal in magnitude and masking is maximal when target and mask are presented simultaneously. Type A masking functions are typically found in noise masking and masking by light paradigms. These two kinds of masking function have been taken as evidence that masks can disrupt visual processing in one of two ways, either by *interruption masking* or by *integration masking* (Kahneman, 1968). In interruption masking, the mask terminates processing of a preceding target before it is complete. In integration masking, the target and mask fuse to form a perceptual composite, whose signal-to-noise ratio is lower than that of the target in isolation. The visual processes assumed to underlie integration masking and interruption masking predict Type A and Type B masking functions, respectively.

For attention researchers, the distinction between different kinds of masking mechanisms is important because links between masking and attention have been found in a number of settings. In metacontrast and object-substitution masking paradigms, the

* Corresponding author. Tel.: +61 03 8344 6343; fax: +61 03 9347 6186.
E-mail address: philipls@unimelb.edu.au (P.L. Smith).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.04.024

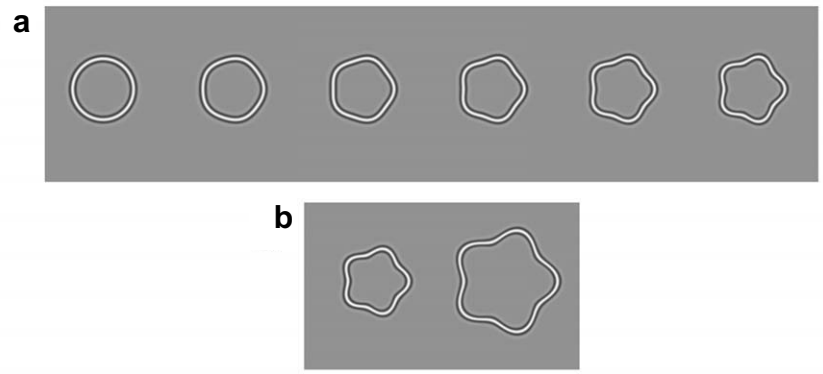

Fig. 1. (a) Example RF patterns. The stimuli are RF(5) patterns with modulation fractions of 0, 0.024, 0.048, 0.072, 0.096, 0.120. (b) RF(5) target stimulus and mask. The modulation of the mask is held fixed while the modulation of the target varies. The target and mask are both presented at the same phase angle, which varies randomly from trial to trial.

magnitude of the masking effect depends on whether or not the stimulus is attended (Enns & Di Lollo, 1997; Ramachandran & Cobb, 1995). Conversely, in spatial cuing and attentional blink paradigms, the magnitude of the attentional effect depends on whether stimuli are masked and on the way they are masked (Enns, 2004; Giesbrecht & Di Lollo, 1998; Lu & Dosher, 1998; Smith, 2000; Smith & Wolfgang, 2007). These findings suggest that studying the relationship between attention and masking may help us understand how attention affects visual processing in a more general way. In this article, we investigate this relationship in an RF masking paradigm.

1.1. The mask-dependent cuing effect in visual signal detection

Our investigation of the link between attention and RF masking grew out of a series of studies by Smith and colleagues on the role of attention in visual signal detection (Gould, Wolfgang, & Smith, 2007; Smith, 1998, 2000; Smith, Ratcliff, & Wolfgang, 2004; Smith & Wolfgang, 2004; Smith, Wolfgang, & Sinclair, 2004; Smith & Wolfgang, 2007). These studies investigated one of the enduring questions in attention, namely, whether detecting a simple, well localized, luminance stimulus in an otherwise empty display benefits from, or requires, attention. The idea that it does not—that detection is, in Neisser's (1967) terminology, a preattentive process—has a long and controversial history that goes back to the first decade of modern attention research, to the pioneering auditory experiments of Cherry (1953) and to the filter theory of Broadbent (1958). Pashler (1998) and Smith (in press) have discussed the theoretical origins of this idea and Palmer, Verghese, and Pavel (2000) have provided a review and an analysis of the recent literature.

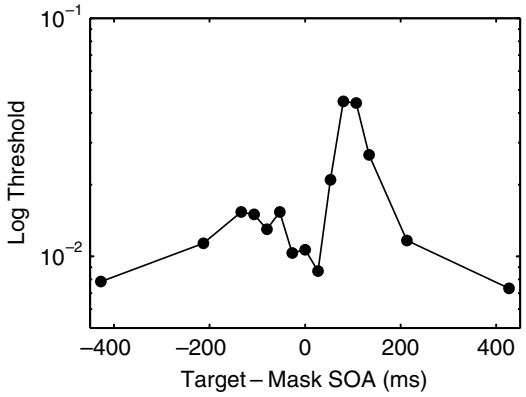

Fig. 2. Type B masking function for RF patterns. The function shows the logarithm of the modulation threshold. There is a relatively small amount of forward masking (negative target-mask SOAs) and a large amount of backward masking (positive cue-target SOAs) that peaks at a critical SOA of around 100 ms. Plot based on data from Habak et al. (2006).

The first study to test the preattention hypothesis using near-threshold stimuli was reported by Bashinski and Bacharach (1980). They found, contrary to the hypothesis, that spatial cues increased detection sensitivity for luminance stimuli. A number of other studies addressed this question using a variety of methods during the following decades, with conflicting results. Studies by Carrasco, Penpeci-Talgar, and Eckstein (2000), Cameron, Tai, and Carrasco (2002), Downing (1988), Hawkins et al. (1990), Luck et al. (1994), Müller and Humphreys (1991), and Smith (1998), reported results consistent with those of Bashinski and Bacharach. These studies all found increased detection sensitivity for attended stimuli. However a second group of studies, by Bonnel, Stein, and Bertucci (1992), Bonnel and Hafter (1998), Brawn and Snowden (2000), Foley and Schwarz (1998), Davis, Kramer, and Graham (1983), Graham, Kramer, and Haber (1985), Lee, Koch, and Braun (1997), Müller and Findlay (1987), Palmer (1994), Palmer, Ames, and Lindsey (1993), and Shaw (1984), found little or no evidence that attention increases detection sensitivity. Several studies in this latter group compared detection and more complex perceptual judgments, such as discrimination or recognition of form (Bonnel & Hafter, 1998; Bonnel et al., 1992; Brawn & Snowden, 2000; Lee et al., 1997; Müller & Findlay, 1987; Palmer, 1994; Palmer et al., 1993; Shaw, 1984). These studies found results consistent with the traditional preattention–attention dichotomy: attention had little or no effect on detection but substantially improved performance for more complex judgments.

Smith (2000) argued that the critical factor distinguishing between the two groups of studies was whether or not backward masks were used to limit the information extracted from the display. He pointed out that, with a few exceptions, studies finding increased sensitivity for attended stimuli limited stimulus information with backward masks. Studies finding no increase in sensitivity limited stimulus information by limiting exposure duration or contrast alone. Discussions of this literature, including an analysis of the exceptional cases, can be found in Gould et al. (2007) and Smith, Wolfgang, et al. (2004). Smith (2000) compared the effects of attention on the detectability of masked and unmasked stimuli in a spatial cuing paradigm and obtained results consistent with this idea. Sensitivity was higher for attended stimuli when stimuli were backwardly masked; when they were unmasked, sensitivity for attended and unattended stimuli did not differ. We refer to this finding as the *mask-dependent cuing effect.*

We have replicated this result in a yes/no task (Smith, 2000; Smith & Wolfgang, 2004), a rating-scale task (Smith, Wolfgang, et al., 2004), and an easy (orthogonal) discrimination task (Smith, Ratcliff, et al., 2004; Smith & Wolfgang, 2007). In this latter task, observers discriminate between vertical and horizontal grating patches in a two-alternative, forced-choice task. Because contrast thresholds in this latter task are indistinguishable from those in yes/no detection (Thomas & Gille, 1979), a number of authors have argued that orthogonal discrimination and yes/no detection are functionally equivalent and interchangeable for the purpose of drawing inferences about attention (Cameron et al., 2002; Lee et al., 1997). The similar patterns of mask dependencies we have found in the two tasks supports this idea.

We have proposed two models of the mask-dependent cuing effect, both of which assume that backward masks act in very simple way, to limit the time during which stimulus information is available to later processing stages. In Coltheart's (1980) terms, masks limit the *informational persistence* of stimuli. We assume that when stimuli are unmasked, informational persistence is long; when they are masked, informational persistence is short because of suppression by the mask. One model is a *rate* or *gain* model (Smith, 2000; Smith & Wolfgang, 2004). It assumes that attention affects the rate at which the outputs of early visual filters are transferred to later processing stages. Carrasco and McElree (2001) reported experimental support for this idea at about the same time. The second model is an *orienting* model (Smith, Ratcliff, et al., 2004). It assumes the act of orienting to a stimulus opens an attentional window, or attentional gate, that transfers the stimulus into visual short-term memory (VSTM) (cf. Reeves & Sperling, 1986). The VSTM trace forms the basis for the observer's perceptual decision. The orienting model assumes that, relative to attended stimuli, the opening of the window for unattended stimuli is delayed. In both models, attention interacts with the differential persistence of masked and unmasked stimuli to predict the mask-dependent cuing effect.

The gain model and the orienting model both assume the interaction between attention and visual masks occurs via interruption masking. This is implicit in the idea that masks act to limit the informational persistence of stimuli. It presupposes that the information in unmasked stimuli is available to later processing stages for a relatively long period but the information in masked stimuli is available for a brief interval only. There are two pieces of empirical evidence that support this view. One is a study by Smith and Wolfgang (2007) that compared cuing effects with simultaneous and delayed noise and pattern masks. The second is a study by Smith and Wolfgang (2004) that compared cuing effects with unmasked, monotopically masked, and dichoptically masked, stimuli.

Smith and Wolfgang (2007) compared the magnitude of the cuing effect with simultaneous masks and delayed (backward) masks, presented at SOAs of 60 or 90 ms. The simultaneous masks were created by presenting the target and mask in interleaved, 10 ms video frames; the delayed masks were created by presenting blank frames between the target and mask frames. In all five of their experiments, they found large and systematic cuing effects with delayed masks and much weaker and less systematic cuing effects with simultaneous masks. As the simultaneous masks should have maximized integration masking and the delayed masks should have maximized interruption masking, the finding that the largest cuing effects were obtained with delayed masks suggests the interaction with attention in detection occurs primarily via an interruption masking mechanism.

Smith and Wolfgang's (2004) investigation of cuing effects with dichoptic and monoptic masks was stimulated by the work of Turvey (1973), who used dichoptic masks to identify central and peripheral components of masking. Turvey showed that the peripheral component depended on the relative energies of the target and mask whereas the central component depended on the target-mask SOA. These results suggest that, in his paradigm, peripheral masking occurred by integration, whereas central masking occurred by interruption. Smith and Wolfgang found that, not only did dichoptic presentation preserve the mask-dependent cuing effect, it also produced a somewhat larger effect than was obtained with monoptic presentation. They argued that interruption masking should be maximized by dichoptic presentation because it prevents the integration of target and mask in primary afferent pathways (cf. Turvey, 1973, for a similar argument). The finding that the cuing effect was also maximized by dichoptic presentation is consistent with the idea that the interaction with attention in this paradigm occurs primarily via an interruption masking mechanism.

1.2. RF masking as a test bed for interruption-mask-dependent cuing

The studies reviewed by Smith (2000) that led to the mask-dependent cuing hypothesis used a variety of different stimulus and mask configurations. Our studies have all used Gaussian-vignetted sinusoidal gratings (Gabor patches) as targets and, most often, high-contrast checkerboards, presented at the same display location, as masks. Our original reason for using checkerboards was to try to create an analogue of a masking-by-structure paradigm, like that used by Turvey (1973), but with grating stimuli. In a typical masking-by-structure paradigm, the stimuli are high contrast, stroke-font, alphanumeric characters and the masks are formed by spatially randomizing features from the target character set. Such masks preserve the spatial frequency composition of the stimuli while randomizing their phases. Consequently, they should stimulate a set of oriented, spatial frequency filters centered on the same frequency as the target but differing in orientation and, possibly also, position in the visual field. Because the contrast energy in checkerboards is concentrated on the diagonals (±45°), a checkerboard mask following a vertically oriented or a horizontally oriented Gabor patch target should likewise stimulate filters that have the same frequency as the target but a different orientation. When followed by a checkerboard mask, grating contrast must be roughly doubled to achieve a level of performance equal to that obtained when no masks are used.

RF masks differ in a fundamental way from the masks we have previously used to study the mask-dependent cuing effect because the contours of an RF mask surround the target but do not cover it. In this, RF masks resemble metacontrast masks (Breitmeyer, 1984) and object-substitution masks (Di Lollo, Enns, & Rensinck, 2000). As we noted earlier, RF masks produce a very distinctive U-shaped, or Type B, masking function, which has often been taken as the signature of an interruption-masking mechanism. Metacontrast masks, in contrast, may produce either a Type A or a Type B masking function, depending on the relative energies of target and mask (Breitmeyer & Ogmen, 2000). If, as we have argued, the mask-dependent cuing effect depends on an interaction between attention and an interruption-masking mechanism, then we would also expect to find a mask-dependent cuing effect with RF-masked targets. We carried out the following experiment as a test of this idea.

2. Methods

2.1. Observers

Three paid undergraduate volunteers, who were naive to the purpose of the experiment, and one of the authors (YL) served as observers. All observers had normal or corrected-to-normal vision. Each observer served in a total of 16 experimental sesssions, each of around one hour duration, preceded by 6–8 practice and calibration sessions.

2.2. Apparatus

Stimuli were presented on a 20 in. Sony G520 monitor at a resolution of 1024×768 pixels. The monitor was driven at a frame rate of 100 Hz by a Cambridge Research Systems ViSaGe visual stimulus generator controlled by a Pentium P4 computer. The display response was linearized (gamma corrected) from measurements made with a Cambridge Research Systems OptiCal photometer. Stimulus presentation and response recording were controlled by software written in C++. Observers performed the task in a dimly-lit laboratory at a viewing distance of 100 cm. with their viewing positions stabilized by a chin rest.

2.3. Stimuli

Following Habak et al. (2006), our stimuli and masks were RF patterns with a radial frequency of 5 (RF5), like those shown in Fig. 1. RF patterns are formed by sinusoidally deforming the mean radius, r_{mean}, of a circle according to the following equation,

$$r(\theta) = r_{\text{mean}}[1 + A\sin(\omega\theta + \phi)],$$

where r and θ are the polar coordinates of the center (i.e., the point of peak luminance) of the contour and ω and ϕ are its radial frequency and phase angle, respectively, expressed in radians. The constant A is the modulation depth, expressed as a proportion of r_{mean}. A modulation depth of zero is a true circle.

The cross-sectional luminance profile of the contour was the fourth derivative of a Gaussian (D4),

$$\text{RF}(d_r) = L_m\left[1 + c\left(1 - 4d_r^2 + \frac{4}{3}d_r^4\right)\exp(-d_r^2)\right].$$

In this equation, L_m is the mean luminance of the display, c is peak stimulus contrast, $d_r = [r(\theta) - r]/\sigma$, is the scaled radial distance from $r(\theta)$, the center of the contour, and σ is the Gaussian space constant. In our experiment, target and mask stimuli had a mean radius, r_{mean}, of 0.60° and 0.92°, respectively, and A, the modulation depth of the mask, was set at 0.12. The mean luminance of the display was 84 cd/m^2, peak stimulus contrast was 0.99, and the Gaussian space constant of the D4 was 0.04°. We used RF5 stimuli as targets and masks in our study and a mask modulation of 0.12 because Habak et al. (2006) obtained a strong Type B masking function using similar stimuli in their study.

We presented targets and masks either simultaneously (SIM), or successively (SUC) at a target-mask SOA of 100 ms. Habak et al. (2006) showed that the peak of the masking function occurred at SOAs of between 80 and 110 ms, so our chosen SOA should have been at or near the peak. In the SIM condition, three 10-ms target frames were interleaved with three mask frames. In the SUC condition, three target frames, interleaved with blank frames, were followed by three mask frames, also interleaved with blanks, at an SOA of 100 ms. The temporal sequence of the stimulus events on an experimental trial is shown in Fig. 3. The inset at the upper right of Fig. 3 shows the sequence of frames used to present the stimulus and mask on SIM and SUC trials.

The attentional cues consisted of four black, right angle markers that identified the corners of a 2.2° square, centered on the target location. The cues were flashed for 60 ms, at a cue-target SOA of 140 ms, and then extinguished. On any trial there were three potential stimulus locations, one cued and two uncued, located at an angular separation of 120° on the circumference of an imaginary 4.5° radius circle, centered on a fixation cross. On any trial, a randomly chosen angle, α ($0 < \alpha \leqslant 360°$) determined the position of the cued location. The possible uncued locations were at $\alpha \pm 120°$. A weakly predictive cuing manipulation was used, similar to that in previous studies of the mask-dependent cuing effect by Smith, Ratcliff, et al. (2004) and Smith and Wolfgang (2007). Stimuli were presented at the cued location on 50% of trials and at each of the uncued locations on 25% of trials, the choice of location on uncued trials being random.

2.4. Design and procedure

Half of the stimuli were true circles and half were RF5 patterns. Five levels of modulation were chosen for SUC stimuli and five for SIM stimuli for each observer during practice to span a range of performance from near chance to near perfect. The modulation fractions for each observer are shown in Table 1. The experiment was run using the method of constant stimuli as a 2 Mask (SUC, SIM) $\times$ 2 Cue (Cued, Miscued) $\times$ 5 Modulation Depth factorial design. Each experimental session comprised 480 trials, yielding a total of 7680 trials per observer.

Each trial began, after a 2.0 s intertrial interval, with presentation of the fixation cross, 1.0 s before the cue. This served both as a warning signal and an instruction to the observer to maintain fixation. Observers judged whether the stimulus was a modulated or an unmodulated circle (i.e., an RF5 or a true circle) and indicated their decision by pressing one of two keys on an infrared response box (CRS CT6). They were instructed to maintain their fixation for the duration of a trial but to try to use the cue to direct their attention to the stimulus location. Because the time from cue onset to stimulus offset (190 ms) was too short to refixate the display (Hallet, 1986), eye movements were not monitored. Observers were given trial-by-trial feedback on their performance auditorily, via distinctive tones, and summary feedback on the visual display at the end of each 20-trial block.

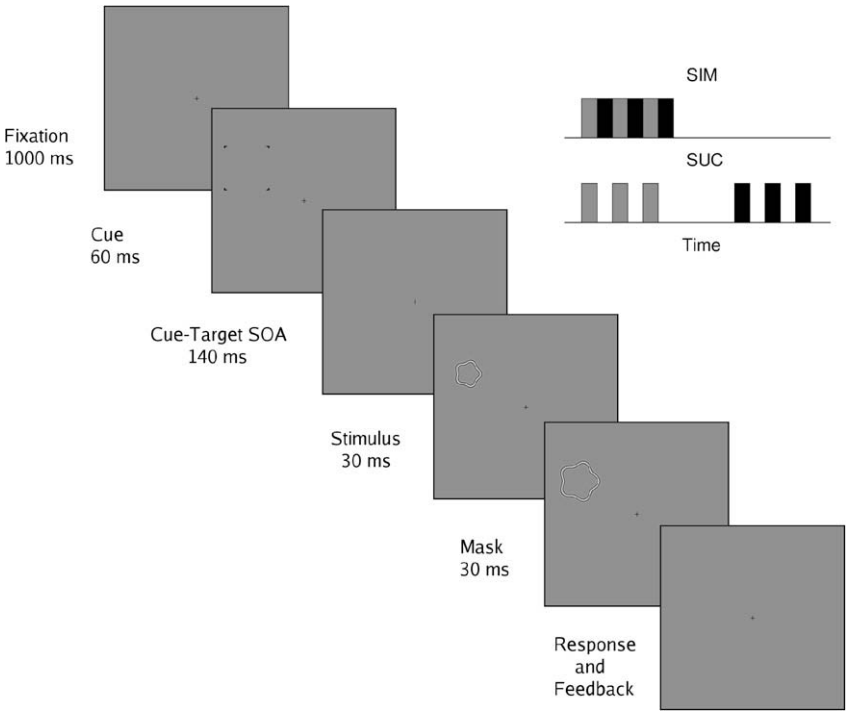

Fig. 3. Event sequence on an experimental trial showing fixation, cue, stimulus, and mask fields. The inset at the upper right shows the event sequence on SIM and SUC trials. Each rectangle in the inset represents a 10 ms frame. Gray rectangles are stimulus frames; black rectangles are mask frames. On SIM trials, three target frames were interleaved with three mask frames. On SUC trials, the three target frames, interleaved with blank frames, were followed after an SOA of 100, by three mask frames, also interleaved with blanks.

Table 1
Modulation fractions for individual observers

Observer	SUC					SIM				
RE	.030	.052	.075	.098	.120	.010	.020	.030	.040	.050
XL	.080	.105	.130	.155	180	.040	.075	.110	.145	.180
FL	.020	032	.045	.058	.070	.010	.012	.015	.018	.020
YL	.020	.028	.035	.042	.050	.010	.018	.025	.032	.040

3. Results and discussion

We converted the proportions of hits and false alarms ("yes" or "signal" responses to modulated and unmodulated stimuli, respectively) to d' sensitivity measures using the formula

$$d' = z(S|s) - z(S|n),$$

where $z(S|s)$ and $z(S|n)$ are the z-scores (inverse Gaussian transforms) of the proportions of hits of and false alarms (i.e., signal responses to signal and noise stimuli, respectively). Plots of d' for cued and miscued stimuli in the SUC and SIM mask conditions are shown for each observer in Fig. 4. To test for cuing effects, we fitted three-parameter Weibull functions of the form

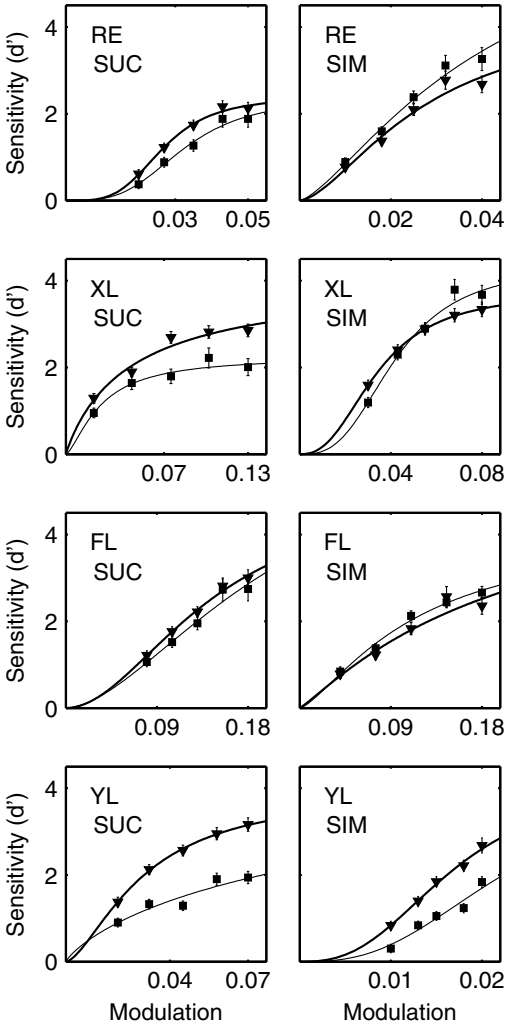

Fig. 4. Psychometric functions for individual observers. The triangles and heavy lines are cued stimuli; the squares and light lines are miscued stimuli. The symbols are the observed data; the continuous curves are fitted Weibull functions. The error bars are asymptotic estimates of the standard error of d'.

$$F(A) = \alpha\left\{1 - \exp\left[-\left(\frac{A}{\beta}\right)^{\gamma}\right]\right\}$$

to the psychometric functions for sensitivity for each observer in the SUC and SIM conditions separately. In this equation, A is the RF modulation depth, α is the amplitude, β is the dispersion (threshold), and γ is the shape (slope). We fitted the psychometric functions by minimizing a chi-square statistic using the Matlab implementation of the Nelder–Mead Simplex algorithm (fminsearch) as described in the Appendix.

For each observer and each mask condition, we compared two models: a single-function model, in which the same Weibull function was fitted to the psychometric functions cued and miscued stimuli, and a two-function model, in which separate Weibull functions were fitted to the psychometric functions for cued and miscued stimuli. We judged the psychometric functions for cued and miscued stimuli to be significantly different if the two-function model provided a significantly better fit to the data than the single-function model. The results of these fits are summarized in Table 2. The fit statistics in the table are the differences in fit between the one-function and the two-function model and are distributed approximately as $\chi^2(3)$ random variables.

For all observers in the SUC condition, sensitivity for cued stimuli was higher than for miscued stimuli. For three of the observers (RE, XL, and YL) the difference was highly significant; for the fourth, (FL) it did not reach significance. These three observers (RE, XL, and YL) also showed significant differences in sensitivity for cued and miscued stimuli in the SIM condition. Two of the observers (RE and XL) showed reversals, with sensitivity to miscued stimuli higher than to cued stimuli, whereas YL again showed higher sensitivity to cued than to miscued stimuli. FL showed a small reversal, but the difference was again not significant. We also fitted a restricted two-function model to the data for FL, in which only the Weibull amplitude parameter was allowed to vary across cue conditions, but the threshold and slope were constrained to be equal. For the restricted model, the difference in sensitivity for FL in the SUC condition was significant, $\chi^2(1) = 5.63$, $p < .05$, as was the reversal in the SIM condition, $\chi^2(1) = 4.10$, $p < .05$.

The pattern of performance in Fig. 4 is strongly reminiscent of that found in previous studies of the mask-dependent cuing effect (Smith, 2000; Smith & Wolfgang, 2004; Smith, Ratcliff, et al., 2004; Smith, Wolfgang, et al., 2004; Smith & Wolfgang, 2007). For the majority of observers, sensitivity to cued stimuli was higher than to miscued stimuli when stimuli were backwardly masked (SUC) and for the majority of observers there was no sensitivity advantage for cued stimuli in the SIM condition. When the restricted rather than the unrestricted model is considered, all observers in the SUC condition showed a sensitivity advantage for cued stimuli, whereas only one observer showed an advantage in the SIM condition. The differences between the two conditions are clearer when we consider averaged effects across observers.

Fig. 5 shows the average psychometric functions for cued and miscued stimuli in the SUC and SIM conditions. The abscissa values in these plots are the average modulation depths at each of the five

Table 2
Cuing effects for individual observers

Observer	SUC	SIM
RE	17.88***	10.18*
XL	37.99***	10.84*
FL	5.95	4.31
YL	158.74***	104.84***

Note. Table entries are $\chi^2(3)$.
* $p < .05$.
*** $p < .001$.

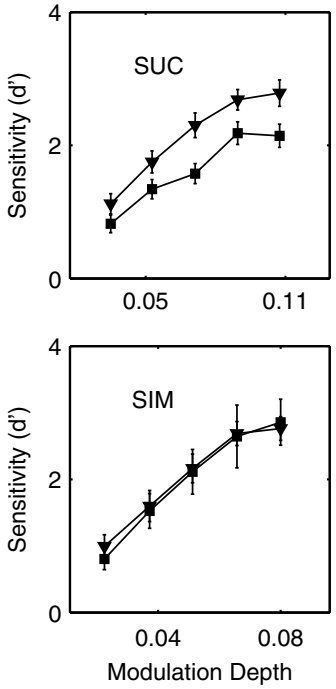

Fig. 5. Sensitivity for SUC and SIM conditions averaged across observers. The triangles are cued stimuli; the squares are miscued stimuli. The error bars are standard errors of the between-subjects means.

levels of modulation for each observer. The averaged effects were tested for significance in a 2 Mask × 2 Cue × 5 Modulation repeated-measures analysis of variance (ANOVA). This analysis yielded a significant main effect for Modulation, $F(4, 12) = 126.47$, $p < .001$, a significant Mask × Cue interaction, $F(1,3) = 17.98$, $p < .05$, and a significant Mask × Cue × Modulation interaction, $F(4,12) = 3.32$, $p < .05$. The main effect of Modulation reflects the increase in d' with increasing depth of RF modulation. The Mask × Cue interaction reflects the fact that mean d' is higher for cued than for miscued stimuli in the SUC condition but does not differ in the SIM condition (the mask-dependent cuing effect). The Mask × Cue × Modulation interaction reflects the fact that the magnitude of the Mask × Cue interaction tends to increase with increasing depth of modulation.

Our experiment has thus replicated the mask-dependent cuing effect with backwardly masked RF patterns, although there were appreciable inter-observer variations. The result is a striking one because the RF masking paradigm is very different to the one used in previous studies of mask-dependent cuing. Whereas those studies used masks presented at the same location as the target, which overwrote the target perceptually, the contours of an RF mask surround the target but do not cover it. Our results thus represent a significant generalization of the mask-dependent cuing effect, and show that the processes underlying the effect operate in more general masking paradigms than the one in which it was first identified. They also add weight to the evidence from Smith and Wolfgang (2004, 2007) that the interaction with attention in the mask-dependent cuing effect occurs via an interruption masking mechanism. As discussed earlier, Habak et al. (2006) showed that RF masking produces a very pronounced Type B masking function, of the kind found by Turvey (1973) in his study of dichoptic, central masking. Such masking functions are usually interpreted as evidence that a trailing mask, presented at a critical SOA, disrupts perceptual processing of the target before it is complete—that is, of an interruption masking mechanism of some kind.

As we remarked above, the pattern of mask dependencies, including the individual differences, is strongly reminiscent of that found in previous studies of the mask-dependent cuing effect. Although these studies have reliably found cuing effects with backwardly masked stimuli, the magnitude of the inter-observer variability has usually been appreciable. Whereas some observers show very large cuing effects others show weak or, sometimes, nonsignificant effects. These individual differences are not apparent in the typical attention experiment in cognitive psychology, in which only averaged effects are reported, but they appear routinely in psychophysical small-*N* designs, in which data are analyzed on an observer-by-observer basis. As yet, there is no good theory that explains why these differences should arise.

The other noteworthy feature of the results in Fig. 4 is the reversal in the cuing effect in the SIM condition. Two of the observers showed significant reversals according to the unrestricted model (separate three-parameter Weibull fits for the cued and miscued condition); three observers showed significant reversals according to the restricted model. These results echo those previously found with unmasked stimuli. Smith (2000), Smith, Wolfgang, et al. (2004), and Smith, Ratcliff, et al. (2004) all found that some observers showed higher sensitivity to miscued than to cued stimuli—an effect they attributed to forward masking of the stimulus by the cue. If correct, this explanation implies that the magnitude of the cuing effect with backwardly masked stimuli will be underestimated, because any forward masking by the cue will occur with masked and unmasked stimuli alike. Any advantage conferred by the cue when stimuli are backwardly masked will have to exceed the masking or inhibition caused by the cue before it can be identified.

One of the observers, YL, showed significant cuing effects in both the SUC and the SIM condition. This result again echoes the findings of Smith and Wolfgang (2007) who compared the magnitude of the cuing effect with simultaneously masked and backwardly masked stimuli. They found that the largest cuing effects were obtained with backward masks, but a proportion of observers also showed significant cuing effects with simultaneous masks, as we have found here. In summary then, both the overall effects and the pattern of individual differences in our study closely parallel previous findings obtained using grating stimuli and masks presented at the same location as the target. These results demonstrate the generality of the effect and show it is not restricted to a particular subset of stimuli and masks.

4. A model of the mask-dependent cuing effect

As noted above, Smith and Wolfgang (2004) and Smith, Ratcliff, et al. (2004) proposed models of the mask-dependent cuing effect. One model is a gain model; the other is an orienting model, but both assume the mask-dependent cuing effect results from an interaction between attention and the differential persistence of masked and unmasked stimuli. More recently, we have developed a general theory that combines elements of both of these models (Smith, 2005, in press). The main elements of the new theory are depicted in Fig. 6. As the figure shows, our new theory seeks to link visual encoding, masking, visual short-term memory (VSTM), and perceptual decision making in a stochastic, dynamic framework. A detailed account of this theory will be presented elsewhere (Smith & Ratcliff, submitted for publication) and we do not attempt to give a full summary of it here. Instead, we use a simplified version of it to show how the mask-dependent cuing effect is obtained as a consequence of two assumptions: (a) that attention affects the rate at which the outputs of early visual filters are transferred to later processing stages, and (b) that backward masks limit the time during which this transfer can occur.

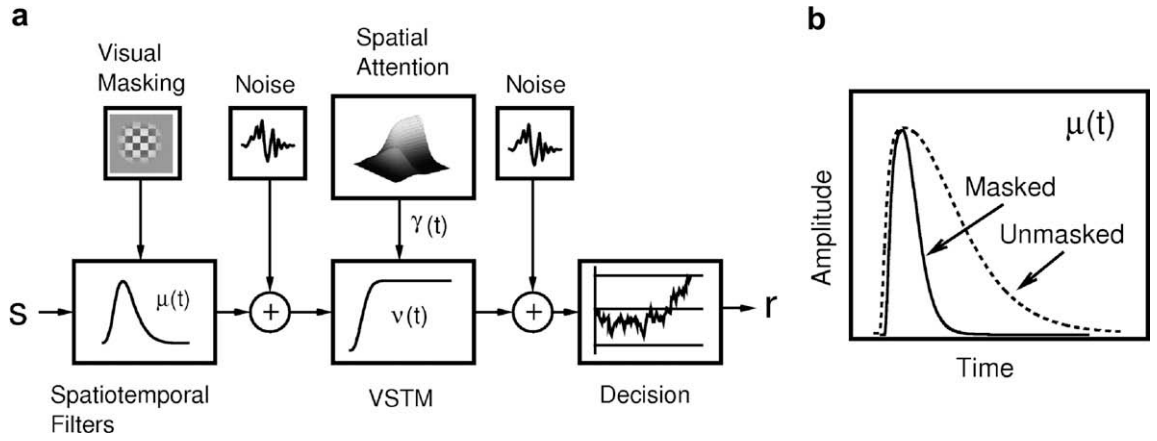

Fig. 6. (a) Stochastic, dynamic model of the mask-dependent cuing effect. The information in the stimulus is represented by a sensory response function, $\mu(t)$. The transient information in the sensory response function is encoded in visual short-term memory (VSTM) under the control of spatial attention. When stimuli are attended, the VSTM trace, $\nu(t)$, grows rapidly to an asymptote; when they are unattended, the trace grows slowly to the same asymptote. The VSTM trace is continuously perturbed by noise. Successive samples of the noise-perturbed VSTM trace are accumulated over time to a criterion to make a decision. The model also has a second source of between-trials noise, which is analogous to the noise in signal detection theory. (b) Example sensory response functions for masked and unmasked stimuli. When stimuli are unmasked, informational persistence is long, allowing the VSTM process to run to asymptote. When stimuli are masked, the mask suppresses the stimulus before the asymptote is reached.

4.1. Visual encoding

As is usual in visual psychophysics, we assume that stimuli are encoded by mechanisms that, to a first approximation, act as linear spatiotemporal filters (Graham, 1989). These filters are low pass in the sense that their response to a brief, rectangular pulse stimulus is a smooth, time-varying function, as shown in Fig. 6. We refer to this function as the *sensory response function* and denote it by $\mu(t)$. We model the sensory response function using cumulative gamma functions of the form

$$\Gamma(t;\beta,n) = 1 - e^{-\beta t}\sum_{j=0}^{n-1}\frac{(\beta t)^j}{j!}.$$

This equation describes the response of a linear filter composed of n cascaded exponential stages, each with rate constant β, to a unit step function. This latter function is zero for t less than zero and unity thereafter. Models based on generalizations of the preceding equation are widely used in studies of visual temporal sensitivity (Watson, 1986).

To characterize the difference in informational persistence of masked and unmasked stimuli, we assume that the onset (rise) and offset (decay) times of the filter may be different. We model this difference using an extended linear filter representation of the form

$$\mu(t) = \Gamma(t;\beta_{on},n)[1 - \Gamma(t-d;\beta_{off},n)].$$

This equation, which was introduced by Smith and Wolfgang (2004), describes the response of a filter with rise constant β_{on} and decay constant β_{off} to a stimulus of duration d. When stimuli are unmasked we assume that $\beta_{off} < \beta_{on}$; when they are masked we assume the converse. Examples of the resulting sensory functions are shown in Fig. 6. The preceding equation may be viewed as a generalization of the symmetrical linear filter model,

$$\mu(t) = \Gamma(t;\beta,n) - \Gamma(t-d;\beta,n),$$

which has been used by many authors (e.g., Busey & Loftus, 1994; Sperling & Weichselgartner, 1995). For $\beta_{off} = \beta_{on}$, Smith and Wolfgang's function is virtually indistinguishable from the symmetrical model for values of β and d of the kind that would be encountered in practice. When $\beta_{off} < \beta_{on}$, $\mu(t)$ describes the response of a filter in which the encoded sensory response is subject to slow iconic decay after stimulus offset; when $\beta_{off} > \beta_{on}$ it describes a sensory response in which the stimulus information is multiplicatively suppressed by the mask.

In adopting the preceding representation of the sensory response function we are attempting only to model the time course of stimulus processing. We do not seek to model the low-level interactions among pattern-sensitive mechanisms in visual cortex that presumably underlie the interruption masking effect (cf. Francis, 1997; Habak et al., 2004; Ogmen, Breitmeyer, & Melvin, 2003). Naturally, a complete theory of the mask-dependent cuing effect would include a computational model of these interactions also. In other, unpublished work, we have replaced the model of the sensory response function described above with one based on Francis's (2003) model of *efficient masking*, which describes the sensory response using differential equations similar to those used to describe the thermodynamics of cooling liquids. When we do so, we obtain the same results as we obtain using the closed form representation above. We use the closed form representation in this article in preference because it is simpler computationally.

4.2. Visual short-term memory

The transient information in the sensory response function is encoded in a more durable form in VSTM under the control of spatial attention. As shown in Fig. 6, in response to a briefly-flashed stimulus the VSTM trace grows smoothly to an asymptote that depends jointly on the contrast, or—in the present application, the modulation depth—of the stimulus and on its duration. We denote the strength of the VSTM trace as a function of time by $\nu(t)$. In the present setting, the VSTM trace represents the strength of the evidence for radial deformation of the circle. The inclusion of an explicit, dynamic model of VSTM is the main difference between our current theory and the earlier model of Smith and Wolfgang (2004). The idea that there exists such a post-iconic, nonverbal stage of VSTM is a widely accepted one, the classic study being that of Phillips (1974).

Our recognition of the need for a VSTM stage in the theory grew out of the work of Ratcliff and Rouder (2000) who studied distributions of response time (RT) in a two-choice discrimination task with brief, backwardly masked stimuli. Because the time needed to make a perceptual decision is usually much longer than the time for which a masked stimulus is present, they argued that observers must have access to a durable representation of the stimulus

during the 300–1000 ms typically needed to make a decision. If this were not the case, the latter part of the decision process, after presentation of the mask, would be driven by noise alone, rather than by stimulus information. Under these circumstances, the predicted RT distributions become highly skewed, particularly at low stimulus discriminabilities, in a way that is not found empirically.

Smith, Ratcliff, et al. (2004) described a simple computational model of VSTM, in which the information in the sensory response function is transferred into VSTM by the act of orienting to the stimulus. However, their model made the somewhat artificial and unsatisfactory assumption that the strength of the VSTM trace depends on the stimulus information falling within a fixed integration window. The primary motivation for our new theory was to provide a more theoretically principled account of the process of VSTM formation.

We model the process of the VSTM trace formation using a *shunting* differential equation. The distinguishing feature of shunting equations is that the information in the stimulus that drives the system—the so-called forcing function—enters into the equation multiplicatively rather than additively, as occurs in the more familiar linear system model. This gives shunting equations nice properties when modeling short-term memory processes, as we discuss below. A number of authors have used shunting equations in models of biological computation in various settings, the most comprehensive analysis being that of Grossberg and colleagues (see Grossberg, 1988, for a summary). The first use of a shunting equation in visual psychophysics was by Sperling and Sondhi (1968) in a model of flicker detection. However, the application most similar to our own is that of Loftus and colleagues (e.g., Busey & Loftus, 1994) who, like us, used a shunting equation to describe the growth of information in VSTM. In their model, the shunting equation is used to model the growth in the proportion of correct responses over time, whereas in our theory it describes the strength of the VSTM trace.

To ensure that the VSTM trace does not saturate at long stimulus exposures, we assume that the trace arises from an opponent process, or excitatory-inhibitory, coding scheme. The simplest shunting equation of this kind, which suffices for the present application, is

$$\frac{dv}{dt} = \Delta_A \mu(t)[\theta - v(t)] - A_{0.5}\mu(t)v(t).$$

The function $\mu(t)$ is the sensory response function, defined previously, which describes the time course of the stimulus. The constants Δ_A and $A_{0.5}$ specify the strength of the excitatory and inhibitory inputs to the VSTM trace, respectively. The excitatory constant Δ_A is an increasing function of the radial deformation of the stimulus while $A_{0.5}$ acts a kind of "inhibitory surround" against which Δ_A is scaled. (The reason for the notation $A_{0.5}$ will become clear in a moment.) The constant θ is the saturation constant for the VSTM trace, which could be set equal to unity without loss of generality.

Two features of the preceding equation deserve comment. One is that when the stimulus is removed, that is, when $\mu(t) = 0$, the derivative dv/dt becomes zero, and the VSTM trace stops changing. This feature of shunting equations provides a natural way to model how a durable memory trace is computed from a briefly presented stimulus. The other feature is that, assuming an initial condition of $v(t) = 0$, the trace always remains bounded on the interval $[0, \theta]$, regardless of the strength of the input. This ensures that the trace cannot grow unboundedly at long stimulus exposure durations.

Using standard techniques for solving first-order differential equations (see Appendix), it is straightforward to show that the solution of the preceding equation is

$$v(t) = \theta\left(\frac{\Delta_A}{A_{0.5} + \Delta_A}\right)\left\{1 - \exp[-(A_{0.5} + \Delta_A)\int_0^t \mu(s)ds]\right\}.$$

This equation shows that the VSTM trace grows exponentially at a rate $A_{0.5} + \Delta_A$ to an asymptote that is proportional to $\Delta_A/(A_{0.5} + \Delta_A)$. The approach to asymptote depends on the area under the sensory response function $\mu(t)$. Viewed as a function of Δ_A, the asymptote is a negatively-accelerating function

$$F(\Delta_A) = \frac{\Delta_A}{A_{0.5+}\Delta_A}$$

that describes how RF modulation is represented perceptually in VSTM. In psychophysical terms, $F(\Delta_A)$ is a transducer function that maps external stimulus magnitudes into internal psychophysical ones. The constant $A_{0.5}$ is a semisaturation constant, which is the modulation at which the function attains half its theoretical maximum of unity. The notation $A_{0.5}$ serves to emphasize this fact.

The exponential form of the preceding VSTM growth function is striking because it closely resembles the function that is used to fit the growth of accuracy in the response-signal paradigm (Carrasco & McElree, 2001; Dosher, 1976, 1979). In the response-signal paradigm, observers are required to respond before an external deadline and a speed–accuracy tradeoff function is generated by varying the deadline systematically. Data from this paradigm are routinely analyzed using an exponential growth equation that is used to characterize asymptotic accuracy, the rate of accuracy growth, and the point at which accuracy first rises above chance. The exponential function is typically used as a convenient way to characterize these various aspects of performance. We have shown here that exponential functions arise naturally as a consequence of a shunting model of VSTM.

4.3. Attention

The VSTM growth equation in the preceding section can be combined with either a gain model or an orienting model of attention (Smith, 2005; Smith & Ratcliff, submitted for publication). Here, we consider a gain model only, because of its simplicity. The gain model assumes that the rate of growth of the VSTM trace depends on whether or not the stimulus is attended (cf. Carrasco & McElree, 2001; Smith, 2000). We can incorporate this idea in the VSTM equation in a very straightfoward way, by introducing an attentional gain constant, denoted, γ. The VSTM growth equation then becomes

$$\frac{dv}{dt} = \gamma\{\Delta_A \mu(t)[\theta - v(t)] - A_{0.5}\mu(t)v(t)\},$$

which has the solution

$$v(t) = \theta\left(\frac{\Delta_A}{A_{0.5} + \Delta_A}\right)\left\{1 - \exp[-\gamma(A_{0.5} + \Delta_A)\int_0^t \mu(s)ds]\right\}.$$

Minimally, we assume that the gain constant has one value, say γ_A, at attended locations and another value, γ_U, at unattended locations, with $\gamma_A > \gamma_U$. Fig. 7 illustrates the most important properties of the resulting VSTM equation.

In the gain-modulated VSTM equation, the final trace strength depends jointly on the gain constant, γ, and the area under the sensory response function, $\mu(t)$. The gain constant depends on whether or not the stimulus is attended and the area under the sensory response depends on whether or not the stimulus is masked. As Fig. 7 shows, when stimuli are unmasked and stimulus persistence is long, the VSTM trace grows rapidly for stimuli at attended locations and slowly for stimuli at unattended locations, but both traces approach their theoretical maximum of $\Delta_A/(A_{0.5} + \Delta_A)$. When stimuli are masked and stimulus persistence is short, the mask suppresses the stimulus before the trace reaches asymptote. Under

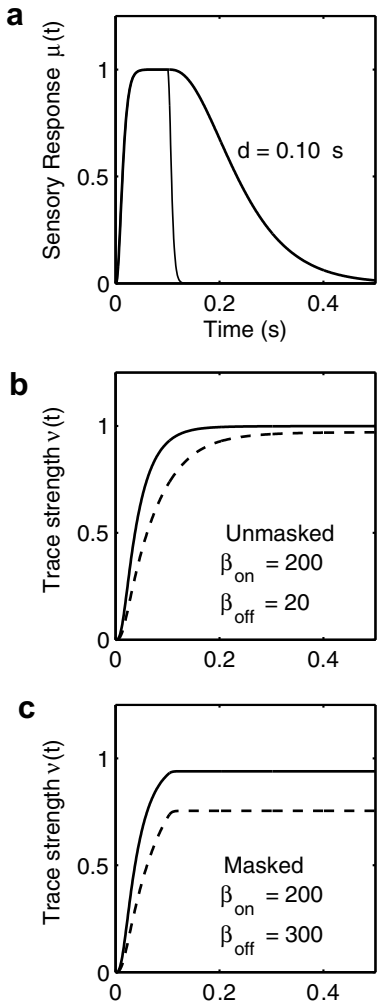

Fig. 7. Predicted VSTM traces for masked and unmasked stimuli, and associated sensory response functions. (a) Sensory response functions, $\mu(t)$, for unmasked (heavy line) and masked (light line) stimuli. The onset rate, β_{on}, was 200 for all stimuli; the offset rate, β_{off}, was 300 for masked stimuli and 20 for unmasked stimuli. (b and c) VSTM traces, $v(t)$, for masked and unmasked stimuli. (b) When stimuli are unmasked, and stimulus persistence is long, the traces for cued stimuli (continuous line) and miscued stimuli (dashed line) grow to the same asymptote. (c) When stimuli are masked, stimulus persistence is short and the VSTM trace does not have time to reach its theoretical maximum. Cued stimuli (continuous line) then have an advantage over miscued stimuli (dashed line) because of their higher rate of VSTM growth. As a result, the traces for cued stimuli will be closer to asymptote when the mask suppresses the stimulus information. The predictions are for a 100 ms stimulus ($d = .10$ s), with no interstimulus interval decay. The gain for cued stimuli was twice the gain for miscued stimuli ($\gamma_A = 2\gamma_U$) and the asymptotic traces strength was set arbitrarily to 1.0.

these circumstances, attended stimuli have an advantage because of their higher rate of VSTM growth. This means that the final trace strength will be greater for attended than for unattended stimuli. This is the mask-dependent cuing effect.

4.4. Decision making

Our theory assumes a sequential-sampling decision process, in which successive samples of the noise-perturbed VSTM trace are accumulated over time until a criterion amount of information needed for a response is attained. In fits of the theory to data we model the decision process either as a Wiener diffusion model (Ratcliff, 1978) or as a dual diffusion model (Ratcliff & Smith, 2004; Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007). In the latter model, the evidence for competing response alternatives is modeled as a pair of parallel, independent diffusion models. As shown in Fig. 6, both models assume there are (at least) two sources of noise in the system: moment-to-moment perturbation of the VSTM trace, whose cumulative effects are modeled as a diffusion process, and between-trial variability in the VSTM trace mean. The latter source of noise is analogous to the noise in signal detection theory.

We assume a sequential-sampling decision process in order to model the complex relationship between accuracy and RT found in spatial cuing tasks (Smith, Ratcliff, et al., 2004; Gould et al., 2007). Sequential-sampling models provide a natural way to model this relationship, both at the level of mean RT and at the level of the RT distributions (Ratcliff & Smith, 2004). In our present application, however, we have only psychophysical data, so we assume a much simpler decision model. Specifically, we assume that d' in any condition is proportional to VSTM trace strength. This is equivalent to assuming a constant variance signal detection model, in which mean signal strength is proportional to the strength of the VSTM trace. It is consistent with the interpretation of d' as an estimate of the transducer function that describes the deterministic portion of the internal representation of a stimulus.

We fitted the VSTM model to the average d' for SIM and SUC conditions in Fig. 4 by weighted least squares, with the reciprocal of the between-subjects standard error of the mean used as the weight for each data point. A version of the fitted model is shown in Fig. 8. As can be seen in the figure, the model nicely reproduces the mask-dependent cuing effect: predicted sensitivity is higher for cued than for miscued stimuli at all modulation depths in the SUC condition, but does not differ in the SIM condition.

The fitted model shown in Fig. 8 had one further elaboration we have not yet described. Rather than assuming the perceptual representation of RF modulation is veridical, we assumed it is subject to an early, nonlinear, power-law transformation, with exponent ρ. This leads to a modified transducer function of the Naka-Rushton form,

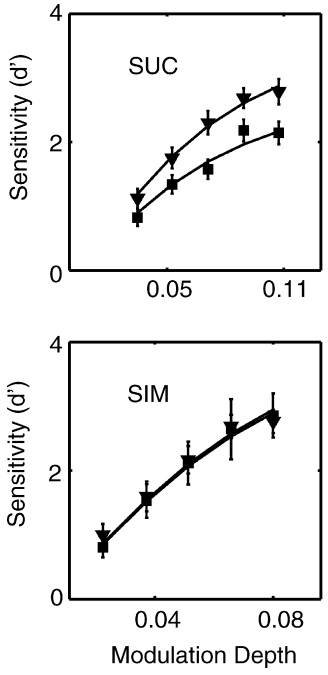

Fig. 8. Fit of the VSTM model to the average data for SUC and SIM trials. The triangles are cued stimuli; the squares are miscued stimuli. The continuous curves are fitted values for the model. The different predictions for SUC and SIM trials were obtained by differences in a single model parameter, the sensory response function offset, β_{off}.

$$F(\Delta_A) = \frac{(\Delta_A)^{\rho}}{A_{0.5}^{\rho} + (\Delta_A)^{\rho}},$$

of the kind that is widely used to model the visual contrast response (e.g., Ling & Carrasco, 2006; Smith, Ratcliff, et al., 2004). With this modification the fit improved from SSE = 5.83 to SSE = 3.82, which is around 50% better than that obtained with $\rho = 1$ (no early nonlinearity). The principal reasons for this improvement was that, with the estimated value of $\rho = 1.51$, the model was better able to capture the bowing over in d' at high modulation depths that appears in Fig. 4.

The estimated parameters of the fitted model are shown in Table 3. These included three sensory response function rate parameters, two attentional gain parameters, and the VSTM parameter, θ, which maps the transducer function to d'. The number of cascaded stages in the sensory filter, n, was fixed at 3. The fit of the model is relatively insensitive to the value of this parameter, as is typically found with linear filter models of visual temporal sensitivity. As can be seen in the Table, when stimuli were unmasked $\beta_{off} < \beta_{on}$, and when they were masked $\beta_{off} > \beta_{on}$, and the attention gain parameter for cued stimuli, γ_A, was roughly double that for miscued stimuli, γ_U.

Although the model in Fig. 8 nicely captures the observed mask-dependent cuing effect, we emphasize that the fit shown in the figure is illustrative only, to show the working of our theory. Our reason for the developing the theory, and much of its power, comes from its ability to describe the relationship between accuracy and the distributions of RT for correct responses and errors in spatial cuing paradigms. Under these circumstances, a more complex model of VSTM is required, in order to capture the way the shapes of the RT distributions change across different experiments. However, a discussion of these issues is outside the scope of this article.

5. Relationship to other work

5.1. External noise exclusion and the perceptual template model

Throughout this article, we have emphasized the role of interruption masking as the basis for the mask-dependent cuing effect. Our finding of a mask-dependent cuing effect in the RF masking paradigm supports this interpretation. However, a different perspective on the relationship between attention and visual masking can be found in the work of Dosher, Lu, and colleagues (Lu & Dosher, 1998; Dosher & Lu, 2000a, 2000b). In a series of articles beginning in the late 1990s, they investigated the role of attention in cued discrimination and recognition tasks using a perceptual template model, an extended signal detection model that incorporates the effects of multiple noise sources, both internal to, and external to the observer. One of their most robust and well-replicated findings is that attentional effects increase in magnitude when stimuli are embedded in a background of external noise. They have attributed this increase to an attention-dependent mechanism of *external noise exclusion*. This mechanism allows the observer to filter out noise at the target location and to obtain a sharper perceptual representation of attended stimuli.

Table 3
Parameters of the VSTM model

Parameter		Value
Sensory response function onset rate	β_{on}	162.9
Sensory response function offset rate (masked)	$\beta_{off}(m)$	375.9
Sensory response function offset rate (unmasked)	$\beta_{off}(u)$	11.6
Attention gain (attended)	γ_A	7.9
Attention gain (unattended)	γ_U	4.2
Naka-Rushton amplitude	θ	4.0
Naka-Rushton exponent	ρ	1.51
Naka-Rushton semisaturation	$A_{0.5}$	0.05

The stimuli in Lu and Dosher's studies of external noise exclusion were presented in alternate video frames, interleaved with frames of uncorrelated Gaussian noise. The noise thus acted as an integration mask, as they themselves have noted, similar to the noise mask used in Turvey's (1973) two-process analysis of central and peripheral masking. Their research can therefore be seen as having shown a relationship between attention and integration masking. Smith and Wolfgang (2007) carried out their study with SIM and SUC masks to try to test whether the mask-dependent cuing effect is another manifestation of Lu and Dosher's external noise exclusion mechanism. To the extent that a mask may be viewed as a source of noise in the display, such an identification is plausible, as Lu, Lesmes, and Dosher (2002) argued.

Smith and Wolfgang (2007) argued that if the mask-dependent cuing effect is the result of an external noise exclusion mechanism, then the cuing effect should follow a Type A masking function. It should be maximal when target and mask are simultaneous and should decrease with increasing temporal separation. If, on the other hand, the cuing effect is the result of an interruption-masking mechanism, it should follow a Type B masking function. It should be maximal when the mask trails the target by a critical SOA and should decrease at shorter or longer SOAs. As we noted previously, the largest cuing effects were obtained in the SUC condition, consistent with an interruption masking account. Nevertheless, a subset of observers in all experiments also showed some evidence of a cuing effect in the SIM condition. This finding contrasts with experiments using unmasked stimuli by Smith (2000), Smith, Ratcliff, et al. (2004), Smith and Wolfgang (2004), and Smith, Wolfgang, et al. (2004), none of which found evidence for a cuing advantage.

Smith and Wolfgang (2007) argued that the pattern of cuing effects in their data was evidence for both kinds of mechanism, a weak, integration-masking-dependent mechanism in the SIM condition and a strong, interruption-masking-dependent mechanism in the SUC condition. They argued that the cuing effect in the SIM condition was probably a manifestation of Lu and Dosher's external noise exclusion mechanism but that the effect in the SUC condition was produced by a mechanism with different dynamics. Smith (in press) suggested that cuing effects of both kinds could be produced by a VSTM model similar to the one described here. He argued that the VSTM model does not require backward masks to predict cuing effects; rather, it will predict a cuing effect whenever the product of attentional gain and the area under the sensory respond function is insufficient to allow the VSTM trace to reach asymptote.

Smith (in press) proposed that one of the effects of embedding stimuli in noise may be to slow the rate at which a perceptual representation of the stimulus is formed. Computationally, the effects of external noise would be modeled as a reduction in β_{on}, the onset rate of the sensory response function. For a stimulus of fixed duration, reducing β_{on} reduces the area under the sensory response function, just as a backward mask does. Cuing effects should therefore be found when stimuli are embedded in external noise, even in the absence of backward masks. Fig. 9 shows how the predictions of the VSTM model change with changes in the size of β_{on}. When β_{on} is large and stimuli are unmasked, the trace grows to its theoretical maximum and no cuing effect occurs. When β_{on} is small and stimuli are unmasked, the area under $\mu(t)$ is insufficient to allow the trace to reach its maximum, and a cuing effect occurs.

This analysis, which attributes the external noise effect to the dynamics of VSTM trace formation, does not contradict the perceptual template account of Lu and Dosher. Rather, it offers a complementary perspective on the same phenomenon, one which stresses the role of temporal processes in the production of cuing effects.

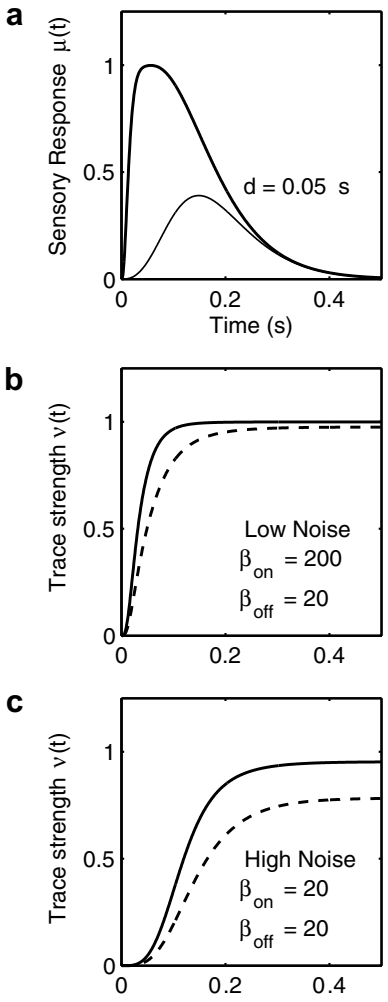

Fig. 9. Predicted VSTM traces, $v(t)$, for stimuli embedded in low and high levels of external noise, and associated sensory response functions, $\mu(t)$. (a) Sensory response functions for low noise (heavy line) and high noise (light line) stimuli. The hypothesized effect of external noise is to reduce β_{on}, the onset rate of the sensory response function, $\mu(t)$, from 200 to 20. The predictions are for a 50 ms stimulus (d = .05 s) with an offset rate of β_{off} = 20 in all conditions. (b and c) VSTM traces for low noise and high noise stimuli. The gain for cued stimuli was twice the gain for miscued stimuli ($\gamma_A = 2\gamma_U$) and the asymptotic strength of $v(t)$ was set arbitrarily to 1.0. The figure shows that cuing effects are predicted with unmasked stimuli when the rate at which stimulus information becomes available perceptually is low.

Whereas Lu and Dosher's perceptual template model assumes that attention allows the observer to filter out noise at the target location to obtain a sharper representation of the signal, the VSTM model assumes the attentional benefit arises because of the higher gain for attended stimuli. Higher gain leads to a stronger final VSTM trace only when the area under the sensory response function is small. We hypothesize that this occurs when stimuli are embedded in noise. The VSTM and perceptual template accounts are compatible theoretically in the sense that both mechanisms could operate within the visual system.

The preceding discussion may help shed light on one of the fundamental questions in the attention literature mentioned earlier, namely, why different tasks benefit from attention to differing extents. Whereas our studies of detection have typically found cues increase sensitivity only with backwardly masked stimuli, other researchers have found cuing effects in a variety of low-level visual tasks in the absence of masks. These include acuity judgments (gap detection in Landolt figures; Yeshurun, Williams, & Carrasco, 2002), orientation discrimination (Cameron et al., 2002), and texture segmentation (Yeshurun & Carrasco, 1998). Smith (in press) argued that such task-dependent attentional effects are consistent with the VSTM model, assuming the information needed to perform fine acuity judgments and difficult discriminations becomes available comparatively slowly. This idea is implied by most multiresolution filter models of the visual system and is well supported empirically (Lupp, Hauske, & Wolf, 1976; Watt, 1987).

As in the preceding discussion of external noise, the key assumption is that β_{on} is smaller for acuity judgments and for difficult discriminations than it is for detection. The effective sensory response function for detection would then resemble the heavy curve in Fig. 9a, whereas the function for discrimination and acuity tasks would resemble the light curve. The VSTM traces for attended and unattended stimuli in discrimination and acuity tasks would then resemble the heavy line and the dashed line in Fig. 9c, respectively. The VSTM trace for cued stimuli would exceed that for miscued stimuli at all times and the difference between them would increase to an asymptote. This is consistent with the speed–accuracy tradeoff functions found experimentally by Carrasco and McElree (2001). When elaborated in this way, the model predicts a Cue × Mask × Task interaction: backward masks will (usually) be needed to obtain a cuing effect in detection but may not be required in acuity judgments or similar tasks. Whether or not a cuing effect will found in a particular task will depend on the duration of the stimulus and the shape of the resulting sensory response function.

Naturally, we do not claim that all attentional phenomena are due to an interaction between the rate of perceptual processing and the rate of VSTM formation. However, the idea has an appealing simplicity and has clarified a confusing and inconsistent part of the attention literature (Smith, 2000). Our strategy has therefore been to see how much can be explained by simple dynamic principles of this kind and only to invoke other mechanisms if and when they appear to be required. Thus, for example, an implication of the preceding analysis is that the effects of external noise should be smaller in detection tasks than in discrimination tasks. A comparison of the results of Smith and Wolfgang (2007) with the studies of Lu and Dosher shows a difference of precisely this kind.

5.2. Attention and other forms of masking

In this article, our primary focus has been attention, and the perspective we have adopted on visual masking is that it is a moderator of the attentional effect. As we noted in the introduction, however, attentional effects have been found in studies in which the primary focus has been to study masking. There the perspective is usually that attention is a moderator of the masking effect. For example, Ramachandran and Cobb (1995) showed that metacontrast masking is affected by attention and Enns and Di Lollo (1997) have shown that object-substitution masking is found only when attention is distributed across multiple items in a visual display. Presumably, the attention research and the masking research are investigating the same, or very similar, phenomena and we assume the same kinds of neural computations underlie the effects in both cases.

Among masking researchers, Francis (2000, 2003; Francis and Cho, 2007), in particular, has studied simple computational models in which the magnitude of masking may vary with attention. We noted earlier that, in unpublished work, we have combined the VSTM model with Francis's (2000, 2003) theory of efficient masking. Efficient masking characterizes the processes underlying Type B masking functions by analogy with the problem of deciding on the optimal time to add cream to cool coffee. Maximum cooling is achieved if the cream is added after the coffee has been allowed to stand, rather than if it is added immediately. This is because the rate of natural cooling is highest when the coffee is hottest.

Consequently, if the coffee is allowed to stand before the cream is added, the amount of natural cooling is maximized. Efficient masking is based on similar principles. Because the natural, iconic decay of a stimulus will be fastest immediately after stimulus offset, masking will be maximal if the mask is delayed. This gives the iconic trace an opportunity to decay before the mask is imposed.

From a theoretical perspective, the VSTM model described here and efficient masking combine together in a very natural way, because both are based on shunting equations. The combination of the two is attractive theoretically because both stages of processing operate on similar computational principles. In practice, however, the simple, closed form representation of the sensory response function we used here and efficient masking lead to virtually identical fits to experimental data. Nevertheless, we think it is likely that, in future, models of attention and models of masking will become increasingly more engaged with one another.

In a recent review of the backward masking literature, Breitmeyer and Ogmen (2000) argued that the simple idea that backward masks interrupt stimulus processing is insufficiently precise theoretically, because it does not identify either the neural substrate of masking or the computational principles that underlie it. Their review suggests that backward masking is not a unitary phenomenon but is, rather, one that encompasses a variety of mechanisms operating at different levels of visual processing. Although we agree with Breitmeyer and Ogmen's position, we also believe that, in many paradigms, the aggregate effect of the action of the these mechanisms is to limit the time for which stimulus information is available to later stages of processing. We use the term "interruption" here to refer to the action of a mechanism or a set of mechanisms that produces such an effect.

In the literature, there exist several physiologically-motivated accounts of masking based on artificial neural networks (see Breitmeyer & Ogmen, 2000, for a review). These include a model developed by Francis (1997) based on Grossberg's (1983) boundary countour system and the RECOD model of Ogmen, Brietmeyer and colleagues (Ogmen et al., 2003). RECOD, in particular, predicts weak forward masking effects and strong backward masking effects that resemble, superficially at least, the effects found in RF masking. Although the details of these models differ, both assume that the effect of a backward mask is to suppress, or to truncate, the visual response to a preceding target. This is the sense in which we use the term "interruption." In our theory, the precise manner in which this inhibition occurs is less important than are its consequences for the subsequent formation of the VSTM trace, on which the perceptual judgment depends.

5.3. Forward masking of the target by the cue

For three observers in the SIM condition in our experiment, sensitivity to miscued stimuli exceeded that to cued stimuli. We noted that this result parallels findings obtained with unmasked stimuli in several earlier studies, and that Smith (2000) attributed it to forward masking of the target by the cue. Luck, Hillyard, Mouloua, and Hawkins (1996) argued similarly. In light of our experimental findings with RF masking, however, this interpretation may need to be re-evaluated. The peripheral cue used in our study is like a forward RF mask in that it is a high-contrast, line-element stimulus that surrounds the target. It is unlike an RF mask in that it is a broken rather than a closed contour that is presented at a more distant spatial location than a typical RF mask. This would lead to the expectation that any forward masking effect of a surrounding peripheral cue should be no stronger, and should probably be weaker, than that of an RF pattern presented at the same SOA. The fact that relatively small forward masking effects are found with RF patterns (Habak et al., 2006) makes us question whether the significant sensitivity reversals in the SIM condition in our experiment could plausibly be attributed to forward masking.

An alternative to the forward masking idea is that the cue may transitorily reduce the saliency of the stimulus, and this may counteract the other benefits of the cue. Elsewhere, in order to fit our VSTM model to RT data, we have assumed that the rate of VSTM formation depends on two factors: the attention gain, and the saliency of the stimulus. Attention gain is a top-down process that depends on the state of the observer, whereas saliency depends on the physical characteristics of the stimulus. We assume that saliency is a measure of the extent to which a stimulus is perceptually distinguishable from its surroundings and that salient stimuli are less susceptible to the effects of miscuing than are less salient stimuli. A plausible, alternative interpretation of the sensitivity reversals in the SIM condition here, and in the unmasked conditions in our previous studies, is that the cue produces a transitory reduction in the saliency of a stimulus presented subsequently at the same location. In the absence of any other benefits conferred by attention, a sensitivity reversal would result. Such a saliency-based explanation differs from the forward masking account, which assumes direct, perceptual interference between the sensory representations of the mask and target.

Yet another possibility is that attention may increase the magnitude of the masking effect, leading to worse rather than better performance at cued locations. This idea is supported by two pieces of evidence. First, Habak et al. (2004, 2006) showed that an RF-modulated mask produces a much larger masking effect than does a control, circular mask. Second, Yeshurun and Carrasco (1998, 1999) showed that attention increases the spatial resolution of vision and that the increase in resolution can either improve or impair performance, depending on the task. As these effects are obtained with or without backward masks (Yeshurun, Montagna, & Carrasco, 2008), they evidently do not rely on interruption-masking-dependent processes of the kind we have identified here. Depending on the particular visual filters used to perform the task, an attention-dependent resolution process like the one identified by Yeshurun and Carrasco would lead to an enhanced representation of curvature (i.e., RF modulation) for both targets and masks. The enhanced representation of targets would lead to an increase in the detectability of RF modulation, but the enhanced representation of masks would also lead to an increase in the masking effect. If the second of these effects were the dominant of the two, a reversal would result. Such reversals would only appear in the SIM condition, in the absence of the other, interruption-masking specific, benefits associated with attention in the SUC condition. We have not yet investigated the implications of either of the preceding two theoretical alternatives.

5.4. The signal detection theory approach to attention

The theory we have described in this article emphasizes the role of attention in the formation of a VSTM representation of the stimulus. The magnitude of the predicted cuing effect depends on the strength of the VSTM traces for attended and unattended stimuli. Our dynamic approach to modeling attention differs in a number of ways from the signal detection approach that has been developed recently by a number of researchers, following the pioneering work of Shaw (1982, 1984). These include Dosher and Lu (2000a, 2000b), Lu and Dosher (1998), Palmer et al. (2000), Eckstein, Thomas, Palmer, and Shimozaki (2000), Smith (1998), and many others. A detailed comparison of the two approaches has been made recently by Smith (in press), but the theoretical issues involved are complex and a full discussion of them is outside the scope of this article. Here we restrict ourselves to commenting on one important point of similarity between the approaches which suggests how the differences between them may be resolved in future.

One of the main insights to have come from the signal detection approach is that the attentional effects in many visual search tasks can be described by some version of a maximum-outputs model. In this model, the observer's decision is based on the properties of the strongest, or most "target like," stimulus in the display (see Palmer et al., 2000, for a review). Typically the way in which the visual system computes the maximum is not specified and is not deemed theoretically important. Rather, the emphasis is on the statistical effects on performance of noise from distractors at nontarget locations in the display (e.g., Baldassi & Burr, 2004). One of the striking successes of the signal detection approach has been to show how a variety of attentional phenomena that had previously been attributed to capacity limitations are predicted by a maximum-outputs model operating on a display containing varying quantities of distractor noise (e.g., Eckstein et al., 2000; Palmer, 1994; Palmer et al., 1993).

We have not considered the theoretical effects of multiple sources of noise in this article because our task used a single, high contrast stimulus in a display with no distractors. Consequently, it is unlikely that noise, whether from distractors or from the uniform field itself, had a significant effect on performance. Factors that would ordinarily be influential when near-threshold stimuli are presented directly against a uniform field—such as the possibility of confusing an empty display location with the target—are unlikely to have been influential here (cf. Gould et al., 2007). Although one can always argue that uniform field noise will have an effect when stimuli are presented under conditions of spatial uncertainty, and that the amount of noise will covary with manipulations of attention, uncertainty cannot explain why cuing effects should vary with the presence or absence of masks or with masks of different kinds. Nevertheless, it remains true that target selection processes are likely to play a significant role in any task in which targets are presented among distractors. It is therefore important to know whether the VSTM model could generalize to multielement displays.

Grossberg and his colleagues have analyzed systems of competitively interacting shunting equations supplemented with nonlinear feedback and have shown they can exhibit winner-take-all dynamics (Grossberg, 1988). When the feedback is faster-than-linear, the activity associated with the strongest input grows to its maximum value while the activity associated with all other inputs is suppressed. Winner-take-all dynamics is of course the neural network counterpart of the maximum-outputs decision rule of signal detection theory. The implication of this analysis is that an extended VSTM model based on a system of coupled shunting equations could exhibit similar target selection properties to those identified in signal detection analyses of visual search. Indeed, Lee, Itti, Koch, and Braun (1999) have proposed a signal detection model of visual attention based on winner-take-all computational principles. The model provides a good account of the effects of attention on spatial frequency discrimination and orientation discrimination of targets embedded in masking noise. Our VSTM model may be seen as an attempt to add an additional, dynamic, dimension to analyses of this kind.

6. Conclusions

In this article, we investigated the effects of attentional cuing in an RF masking paradigm. We showed that a mask-dependent cuing effect, similar to that previously found with pattern-masked, sinusoidal grating stimuli, is also found with RF masking. Little or no cuing effect is found when RF patterns are masked with simultaneous masks, but large and systematic effects are found when they are masked with backward masks. The strong Type B masking functions found with RF patterns is evidence that they mask by a process of interruption masking, and is consistent with the idea that the mask-dependent cuing effect occurs when the mask interrupts the transfer of stimulus information from early visual filters to later processing stages. We presented a computational model of the mask-dependent cuing effect, in which masks affect the informational persistence of stimuli and cues affect the rate at which stimulus information is transferred to VSTM. This model nicely captured the interaction between cues and visual masks in our data. Our findings show that the mask-dependent cuing effect occurs in settings other than the one in which it was originally identified, and suggest it may reflect fundamental, underlying properties of attentional processes.

More generally, they show that masking paradigms can be powerful tools for studying the temporal dynamics of attention.

Acknowledgments

The research in this article was supported by Australian Research Council Discovery Grant DP0209249. The experimental data were collected by Young-Eun Lee for an Honours Thesis in Psychology at the University of Melbourne.

Appendix. Model-fitting procedures

To test for significant cuing effects in d' as a function of RF modulation, A, we fitted Weibull functions of the form

$$F(A) = \alpha\left\{1 - \exp\left[-\left(\frac{A}{\beta}\right)^{\gamma}\right]\right\}$$

to the d' values for each observer in the SIM and SUC conditions by minimizing the chi-square statistic

$$\chi^2 = \sum_i \frac{[d'_i(A) - F_i(A)]^2}{\mathrm{var}[d'_i(A)]},$$

where

$$\mathrm{var}(d') = \frac{P(S|s)P(N|s)}{n_s\phi^2[z(S|s)]} + \frac{P(S|n)P(N|n)}{n_n\phi^2[z(S|n)]}$$

is the asymptotic variance estimate of Gourevitch and Galanter (1967). In this equation, n_s and n_n are the number of signal (nonzero modulation) and noise (zero modulation) stimuli in each condition and $\phi(.)$ is the standard normal density function evaluated at the specified abscissa. The other quantities are as defined in the text. We compared the fit of an unrestricted model, $\{\alpha_C, \beta_C, \gamma_C, \alpha_U, \beta_U, \gamma_U\}$, to a null model $\{\alpha., \beta., \gamma.,\}$ where the subscripts C and U denote "cued" and "uncued", respectively. For FL, we also compared a restricted model $\{\alpha_C, \alpha_U, \beta., \gamma.\}$ to the null model. We tested the improvements in fit produced by the unrestricted and restricted models as chi-square statistics with 3 and 1 degree of freedom, respectively.

Solution of the VSTM shunting equation

To solve the VSTM shunting equation we rewrite it as

$$\frac{dv}{dt} + (\Delta_A + A_{0.5})v(t)\mu(t) = \theta\Delta_A\mu(t)$$

and introduce the integrating factor $\exp[(\Delta_A + A_{0.5})\int^t \mu(s)ds]$. Multiplying both sides of the equation by the integrating factor yields

$$\exp\left[(\Delta_A + A_{0.5})\int^t \mu(s)ds\right]\left[\frac{dv}{dt} + (\Delta_A + A_{0.5})v(t)\mu(t)\right] = \theta\Delta_A\mu(t)\exp\left[(\Delta_A + A_{0.5})\int^t \mu(s)ds\right].$$

By the chain rule, the left hand side of this equation can be recognized as an exact differential,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{v(t)\exp\left[(\Delta_A+A_{0.5})\int^t\mu(s)\mathrm{d}s\right]\right\} = \theta\Delta_A\mu(t)\exp\left[(\Delta_A+A_{0.5})\int^t\mu(s)\mathrm{d}s\right],$$

where $\mathrm{d}/\mathrm{d}t$ denotes the time derivative of the expression in braces. Integrating both sides of the preceding equation with respect to time gives

$$v(t)\exp\left[(\Delta_A+A_{0.5})\int^t\mu(s)\mathrm{d}s\right] = \theta\Delta_A\int^t\mu(s)\exp\left[(\Delta_A+A_{0.5})\int^s\mu(z)\mathrm{d}z\right]\mathrm{d}s$$

or

$$v(t)\exp\left[(\Delta_A+A_{0.5})\int^t\mu(s)\mathrm{d}s\right] = \theta\left(\frac{\Delta_A}{\Delta_A+A_{0.5}}\right)\int^t(\Delta_A+A_{0.5})\mu(s) \times\exp\left[(\Delta_A+A_{0.5})\int^s\mu(z)\mathrm{d}z\right]\mathrm{d}s.$$

In this form, the outer integral on the right hand side is exact. After evaluating this integral the equation reduces to

$$v(t)\exp\left[(\Delta_A+A_{0.5})\int^t\mu(s)\mathrm{d}s\right] = \theta\left(\frac{\Delta_A}{\Delta_A+A_{0.5}}\right)\left\{\exp\left[(\Delta_A+A_{0.5})\int^t\mu(s)\mathrm{d}s\right]+K\right\}.$$

With an initial condition of $v(0)=0$, we obtain

$$K = -\theta\frac{\Delta_A}{\Delta_A+A_{0.5}},$$

yielding

$$v(t)\exp\left[(\Delta_A+A_{0.5})\int_0^t\mu(s)\mathrm{d}s\right] = \theta\left(\frac{\Delta_A}{\Delta_A+A_{0.5}}\right)\left\{\exp\left[(\Delta_A+A_{0.5})\int_0^t\mu(s)\mathrm{d}s\right]-1\right\}.$$

Multiplying both sides by $\exp[-(\Delta_A+A_{0.5})\int_0^t\mu(s)\mathrm{d}s]$ we obtain the solution

$$v(t) = \theta\left(\frac{\Delta_A}{\Delta_A+A_{0.5}}\right)\left\{1-\exp\left[-(\Delta_A+A_{0.5})\int_0^t\mu(s)\mathrm{d}s\right]\right\},$$

as given in the text.

References

Baldassi, S., & Burr, D. C. (2004). "Pop-out" of targets modulated in luminance or colour: the effects of intrinsic and extrinsic uncertainty. *Vision Research, 44*, 1227–1233.

Bashinski, H. S., & Bacharach, V. R. (1980). Enhancements of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception & Psychophysics, 28*, 241–248.

Bell, J., Badcock, D. R., Wilson, H. R., & Wilkinson, F. (2007). Detection of shape in radial frequency contours: Independence of local and global form information. *Vision Research, 47*, 1518–1522.

Bonnel, A-M. A.-M., & Hafter, E. R. (1998). Dividing attention between simultaneous auditory and visual signals. *Perception & Psychophysics, 60*, 113–124.

Bonnel, A-M. A.-M., Stein, J-F. J.-F., & Bertucci, P. (1992). Does attention modulate the perception of luminance changes? *Quarterly Journal of Experimental Psychology, A, 44*, 601–626.

Brawn, P. M., & Snowden, R. J. (2000). Attention to overlapping objects: Detection and discrimination of luminance changes. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 342–358.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. Oxford, U.K.: Clarendon Press.

Breitmeyer, B. D., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62*, 1572–1595.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon.

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review, 101*, 446–469.

Cameron, E. L., Tai, J. C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research, 42*, 949–967.

Carrasco, M., & McElree, B. (2001). Covert attention speeds the accrual of visual information. *Proceedings of the National Academy of Science of the United States of America, 98*, 5363–5367.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research, 40*, 1203–1215.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America, 25*, 975–979.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27*, 183–228.

Davis, E. T., Kramer, P., & Graham, N. (1983). Uncertainty about spatial frequency, spatial position, or contrast of visual patterns. *Perception & Psychophysics, 33*, 20–28.

Di Lollo, V., Enns, J. T., & Rensinck, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*, 481–507.

Dosher, B. A. (1976). The retrieval of sentences from memory: A speed–accuracy study. *Cognitive Psychology, 8*, 291–310.

Dosher, B. A. (1979). Empirical approaches to information processing: Speed–accuracy tradeoff functions or reaction time. *Acta Psychologica, 43*, 347–359.

Dosher, B. A., & Lu, Z.-L. (2000a). Mechanisms of perceptual attention in precuing of location. *Vision Research, 40*, 1269–1292.

Dosher, B. A., & Lu, Z.-L. (2000b). Noise exclusion in spatial attention. *Psychological Science, 11*, 139–146.

Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 188–202.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics, 62*, 425–451.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research, 44*, 1321–1331.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8*, 135–139.

Foley, J. M., & Schwarz, W. (1998). Spatial attention: Effect of position uncertainty and number of distractor patterns on the threshold-versus-contrast function for contrast discrimination. *Journal of the Optical Society of America, A, 15*, 1036–1047.

Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review, 104*, 572–594.

Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychological Review, 107*, 768–785.

Francis, G. (2003). Developing a new quantitative account of backward masking. *Cognitive Psychology, 46*, 198–226.

Francis, G., & Cho, Y. S. (2007). Testing models of object substitution with backward masking. *Perception & Psychophysics, 69*, 263–275.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1454–1466.

Gould, I. C., Wolfgang, B. J., & Smith, P. L. (2007). Spatial uncertainty explains exogenous and endogenous attentional cuing effects in visual signal detection. Journal of Vision, 7(13): 4, 1–17.

Gourevitch, V., & Galanter, E. (1967). A significance test for one-parameter isosensitivity functions. *Psychometrika, 32*, 25–33.

Graham, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Graham, N., Kramer, P., & Haber, N. (1985). Attending to the spatial frequency and spatial position of near-threshold visual patterns. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 269–284). Hillsdale, NJ: Erlbaum.

Grossberg, S. (1983). The quantized geometry of visual space: The coherent computation of depth, form, and lightness. *Behavioral and Brain Sciences, 6*, 625–657.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks, 1*, 17–61.

Habak, C., Wilkinson, F., & Wilson, H. R. (2006). Dynamics of shape interaction in human vision. *Vision Research, 46*, 4305–4320.

Habak, C., Wilkinson, F., Zakar, B., & Wilson, H. R. (2004). Population coding for complex shapes in human vision. *Vision Research, 44*, 2815–2823.

Hallet, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. I* (pp. 10.1–10.11). New York: Wiley.

Hawkins, H. L., Hillyard, S. A., Luck, S. J., Mouloua, M., Downing, C. J., & Woodward, D. P. (1990). Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 802–811.

Kahneman, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin, 70*, 404–425.

Lee, D. K., Koch, C., & Braun, J. (1997). Spatial vision thresholds in the near absence of attention. *Vision Research, 37*, 2409–2418.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience, 2*, 375–381.

Ling, S., & Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research, 46*, 1210–1220.

Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*, 1183–1198.
Lu, Z.-L., Lesmes, L. A., & Dosher, B. A. (2002). Spatial attention excludes noise at the target location. *Journal of Vision, 2*, 312–323.
Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cuing on luminance detectability: Psychophysical and electrophysical evidence. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 887–904.
Luck, S. J., Hillyard, S. A., Mouloua, M., & Hawkins, H. L. (1996). Mechanisms of visual-spatial attention: Resource allocation or uncertain reduction. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 725–737.
Lupp, U., Hauske, G., & Wolf, W. (1976). Perceptual latency to sinusoidal gratings. *Vision Research, 16*, 969–972.
Müller, H. J., & Findlay, J. M. (1987). Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophysics, 42*, 383–399.
Müller, H. J., & Humphreys, G. W. (1991). Luminance-increment detection: Capacity-limited or not? *Journal of Experimental Psychology: Human Perception and Performance, 17*, 107–124.
Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
Ogmen, H., Breitmeyer, B., & Melvin, R. (2003). The what and where in visual masking. *Vision Research, 43*, 1137–1350.
Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research, 34*, 1703–1721.
Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 108–130.
Palmer, J. A., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*, 1227–1268.
Pashler, H. E. (1998). *The psychology of attention*. Cambridge; MA: MIT Press.
Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*, 283–290.
Ramachandran, V. S., & Cobb, S. (1995). Visual attention modulates metacontrast masking. *Nature, 373*, 66–68.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.
Ratcliff, R., & Rouder, J. (2000). A diffusion model account of masking in letter identification. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 127–140.
Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*, 333–367.
Ratcliff, R., Hasegawa, Y. T., Hasegawa, R. P., Smith, P. L., & Segraves, M. (2007). Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *Journal of Neurophysiology, 97*, 1756–1774.
Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review, 93*, 180–206.
Shaw, M. I. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology, 14*, 353–409.
Shaw, M. I. (1984). Division of attention among spatial locations: A fundamental differences between detection of letters and the detection of luminance increments. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance. X* (pp. 109–121). Hillsdale, NJ: Erlbaum.
Smith, P. L. (1998). Attention and luminance detection: A quantitative analysis. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1401–1420.
Smith, P. L. (2000). Attention and luminance detection: Effects of cues, masks, and pedestals. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1401–1420.
Smith, P. L. (2005). An integrated theory of attention and decision making in visual signal detection. Paper presented at the Cold Spring Harbor Workshop on the Psychobiology of Decision Making. Cold Spring Harbor, NY.
Smith, P. L. (in press). Attention and the detection of weak visual signals. In Coltheart, V. (Ed.), Tutorials in visual cognition. Hove, Sussex: Psychology Press.
Smith, P. L., & Ratcliff, R. (submitted for publication). An integrated theory of attention and decision making in visual signal detection.
Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision Research, 44*, 1297–1320.
Smith, P. L., & Wolfgang, B. J. (2004). The attentional dynamics of masked detection. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 119–136.
Smith, P. L., & Wolfgang, B. J. (2007). Attentional mechanisms in visual signal detection: The effects of simultaneous and delayed noise and pattern masks. *Perception & Psychophysics, 69*, 1093–1104.
Smith, P. L., Wolfgang, B. J., & Sinclair, A. (2004). Mask-dependent attentional cuing effects in visual signal detection: The psychometric function for contrast. *Perception and Psychophysics, 66*, 1056–1075.
Sperling, G., & Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America, 58*, 1133–1145.
Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review, 102*, 503–532.
Thomas, J. P., & Gille, J. (1979). Bandwidths of orientation channels in human vision. *Journal of the Optical Society of America, 69*, 652–660.
Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 80*, 1–52.
Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs performance by enhancing spatial resolution. *Nature, 396*, 72–75.
Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research, 23*, 983–989.
Yeshurun, Y., Montagna, B., & Carrasco, M. (2008). On the flexibility of sustained attention and its effects on a texture segmentation task. *Vision Research, 48*, 78–93.
Yeshurun, Y., Williams, P. E., & Carrasco, M. (2002). Covert attention increases spatial attention with or without masks: Support for signal enhancement. *Journal of Vision, 2*, 467–479.
Watson, A. B. (1986). Temporal sensitivity. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.). *Handbook of perception and performance* (Vol. 1, pp. 6.1–6.8). New York: Wiley.
Watt, R. J. (1987). Scanning from course to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of Optical Society of America, A, 4*, 2006–2021.